P-values again!

The p-value is the probability under null of seeing as or more extreme data.

If the p-value is small

=> we are already at the extreme

=> unlikely for null distribution to have generated test statistic

=> reject null (if small enough)

Our arbitrary rules

$p > 0.05$ (5%)

Fail to reject H_0.

(Don't say accept)

$0.01 < p \leq 0.05$

reject, "statistically significant"

Big difference between statistical significance and practically significance.

$p \leq 0.01$

"highly significant"

$p \approx 0 \quad (|z| > 4)$

---

p-value is (sort of)

$P(\text{Data} \mid H_0)$

NOT

~~$P(H_0 \mid \text{Data})$~~

What does "under the null hypothesis" mean?

Based off lecture example

63 stroke victims in sample

83% improved w/ med

in population

53% improvement in general

H_0: difference due to chance

H_1: difference not due to chance

H_0 is equivalent to saying that the true recovery is 53%

H_1 is equivalent to saying true recovery is $>53\%$.

Assuring H_0

tells us that the box

53	47

under null we know

$EV_{\%} = 53\%$

$SD_{box} = \sqrt{0.53 \cdot 47} \approx 0.5$

(not bootstrapping)

$$SE_{\%} = \frac{SD \cdot 100\%}{\sqrt{\#draws}} = \frac{0.5}{\sqrt{63}} \cdot 100\% \approx 6.4\%$$

$$z = \frac{83\% - 53\%}{6.4\%} \approx 4.7\%$$

$1 - \text{pnorm}(4.7)$

$= 1.300807 \cdot 10^{-6}$

$\approx 0.$

reject